Nontarget Screening Reveals Time Trends of Polar Micropollutants in a Riverbank Filtration System


Published in: Environmental Science and Technology

DOI: 10.1021/acs.est.9b01750

Link to publication

Creative Commons License (see https://creativecommons.org/use-remix/cc-licenses):
CC BY-NC-ND

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (http://dare.uva.nl)
Nontarget Screening Reveals Time Trends of Polar Micropollutants in a Riverbank Filtration System

Vittorio Albergamo,* Jennifer E. Schollée,† Emma L. Schymanski,* Rick Helmus,** Harrie Timmer,∥ Juliane Hollender,§ and Pim de Voogt†

†Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, The Netherlands
‡Eawag, Swiss Federal Institute of Aquatic Science and Technology, Überlandstrasse 133, 8600 Dübendorf, Switzerland
§Luxembourg Centre for Systems Biomedicine, University of Luxembourg, House of Biomedicine II 6, avenue du Swing, L-4367 Belvaux, Luxembourg
∥Oasen, Nieuwe Gouwe O.Z 3, 2801 SB Gouda, The Netherlands
**Institute of Biogeochemistry and Pollutant Dynamics, ETH Zürich, Universitätstrasse 16, 8092 Zürich, Switzerland
††KWR Watercycle Research Institute, Groningenhaven 7, 3430 BB, Nieuwegein, The Netherlands

ABSTRACT: The historic emissions of polar micropollutants in a natural drinking water source were investigated by nontarget screening with high-resolution mass spectrometry and open cheminformatics tools. The study area consisted of a riverbank filtration transect fed by the river Lek, a branch of the lower Rhine, and exhibiting up to 60-year travel time. More than 18,000 profiles were detected. Hierarchical clustering revealed that 43% of the 15 most populated clusters were characterized by intensity trends with maxima in the 1990s, reflecting intensified human activities, wastewater treatment plant upgrades and regulation in the Rhine riparian countries. Tentative structure annotation was performed using automated in silico fragmentation. Candidate structures retrieved from ChemSpider were scored based on the fit of the in silico fragments to the experimental tandem mass spectra, similarity to openly accessible accurate mass spectra, associated metadata, and presence in a suspect list. Sixty-seven unique structures (72 over both ionization modes) were tentatively identified, 25 of which were confirmed and included contaminants so far unknown to occur in bank filtrate or in natural waters at all, such as tetramethylsulfamide. This study demonstrates that many classes of hydrophilic organics enter riverbank filtration systems, persisting and migrating for decades if biogeochemical conditions are stable.

INTRODUCTION

Thousands of anthropogenic chemicals are released into the aquatic environment via wastewater treatment plant (WWTP) effluents, runoffs, and accidental spills. Transformation products (TPs) formed during water treatment and under environmental conditions increase the complexity of the chemical mixtures that occur in the environment. Freshwater systems are particularly vulnerable to contamination by polar organic micropollutants (MPs) exhibiting low or negative pH-adjusted log distribution coefficients (logD) as they preferentially partition into the water phase. When persistent, polar MPs can migrate indefinitely throughout the water cycle and reach drinking water sources. In Europe, riverbank filtration (RBF) is a common drinking water pretreatment with potential to remove dissolved MPs mainly by sorption and biodegradation as surface water infiltrates through the hyporheic zone. Sorption to organic matter can delay the transport of neutral and moderately hydrophobic MPs (logD > 3) in RBF systems. The ion-exchange capacity of soils can result in the retention of cationic MPs, but it is not effective on anionic MPs. Biodegradation is favored by a redox potential gradient and long travel time, as they result in greater biodiversity of microbial communities and longer time for adaptation.

Liquid chromatography coupled to high-resolution tandem mass spectrometry (LC-HRMS/MS) is the preferred system to analyze most polar MPs in aqueous matrices. The capability of recent mass analyzers to achieve sensitive detection with high resolving power (>20,000) and high mass accuracy (<5 ppm) is pivotal to tentatively identify unknown ions via accurate
mass spectra without the use of reference standards. In environmental research, these approaches are known as suspect screening and nontarget screening (NTS). Suspect screening and NTS are increasingly being applied to environmental samples and are gradually becoming harmonized. Suspect screening aims at identifying pollutants expected in a sample. Commonly, HRMS1 data (mass-to-charge ratios of ionized analytes) is searched for masses of interest suspected to occur (e.g., for study-specific reasons), typically included in a suspect list, and then accompanying isotopic (and adduct) peaks, HRMS2 spectra (fragment ions), and retention time ($t_R$) are used to support the identification and confirmation of suspect hits. Initiatives to improve screening efforts include platforms to share suspect lists, e.g., the NORMAN Suspect List Exchange and the U.S. Environmental Protection Agency (EPA) CompTox Dashboard, and openly accessible accurate mass spectral libraries, e.g., MassBank and MassBank of North America (MoNA).

In contrast to suspect screening, NTS aims at identifying compounds without searching for certain masses/substances up front, but rather using the measured data to reveal the masses of interest. Since thousands of ions are acquired indiscriminately in HRMS1 full scans, a prioritization strategy is required to select masses of interest. Tentative structures are generally assigned based on candidate searching in chemical databases, e.g., PubChem and ChemSpider, often using HRMS2 spectra. State-of-the-art NTS benefits from the increased availability of computational tools for prioritization, e.g., statistical analysis methods, and cheminformatics tools for high-throughput structure annotation, e.g., in silico fragmentation querying openly accessible chemical databases and accurate mass spectral libraries. An overview on state-of-the-art cheminformatics tools for structure annotation can be found in the literature.

In this study, we investigated a natural drinking water source consisting of a riverbank filtrate originated from the Lek, a branch of the river Rhine in The Netherlands. Bank filtration at this site exhibits up to 60-year travel time from the riverbank to the furthest of a series of wells built by a drinking water utility. This site can be regarded as a hydrogeological archive, where persistent anthropogenic chemicals from the "post-1950s acceleration" to the present are preserved. Our goal was to detect major pollution trends across the bank filtration transect and characterize the identities of mobile MPs by applying state-of-the art nontarget screening. To the best of our knowledge, no previous studies have attempted to investigate a time series of nontarget polar contaminants in a natural bank filtrate with such an extended travel time. Exposure to over a half-century of anthropogenic emissions from intensified industrial and agricultural activities followed by mitigation measures such as wastewater treatment upgrades in the 1990s make this bank filtration system a unique location to detect contamination time series and investigate persistent and mobile MPs in the aquatic environment using nontarget analytical approaches. The occurrence of these chemicals is rationalized and their emission sources discussed. Compounds not previously known to occur in bank filtrate were identified, including chemicals that were not known to occur in the water cycle at all.

### MATERIALS AND METHODS

#### Standards and Reagents
Details on the analytical standards and reagents used for this study are included in the Supporting Information (SI) section S-1.

#### Sampling Site and Sample Collection
Anaerobic bank filtrate with residence times from 1 to 60 years was abstracted from a series of wells located in two adjacent production fields, named Schuwacht and Tiendweg, fed by the river Lek, a branch of the Rhine (Figure 1). The traveling time at each well was previously investigated by the drinking water utility Oasen (Gouda, The Netherlands) by means of isotopic age dating and hydrogeological modeling. The well fields are located in the municipality of Krimpen aan de Lek, South Holland, The Netherlands, and provide raw water for approximately 7000 m$^3$ of drinking water per day. Further details on the riverbank filtration site including a map of the groundwater flow lines with modeled travel time are given in SI section S-2.

Bank filtrate samples from nine wells of increasing travel time ($n = 3$) were collected in 5 L polypropylene bottles from sampling faucets built on each well and immediately transported to the University of Amsterdam, where they were kept in the dark at 2 °C and extracted next day.

#### Solid-Phase Extraction
The samples were allowed to reach room temperature, then 100 mL was transferred to 250 mL polypropylene bottles and spiked with 100 μL of a mix of 128 isotope-labeled internal standards (IS) available at Eawag at concentration of 1 ng/μL. This resulted in IS concentrations of 1 μg/L in the samples before enrichment by solid-phase extraction (SPE). An offline extraction protocol relying on hydrophilic–lipophilic balance (HLB) sorbent with Oasis cartridges by Waters (Etten-Leur, The Netherlands) was adapted from a procedure described elsewhere for the enrichment of moderately hydrophobic and polar organics. The adjustments to the extraction protocol were the spike volume and concentrations of IS and the final concentration step, which in the present study resulted in an enrichment factor of 200 as the extracts were diluted 5-fold with deionized water prior LC-HRMS analysis.

#### LC-HRMS Analysis
The samples were analyzed at Eawag using LC-HRMS. A high-performance liquid chromatography system (HPLC) consisting of a PAL Autosampler (CTC...
Separate analysis runs were conducted for positive and resolving power of 17,500 (more details in SI section S-3.1). Wormer, The Netherlands). Further details on this system are HRMS equipped with an ESI source (Bruker Daltonics, Den Bosch, The Netherlands) coupled to a maXis 4G q-ToF/

sample extracts were analyzed with an ultra-high-performance University of Amsterdam. Reference standard materials and ranging from extraction and chromatographic methods. HRMS detection not be sufficiently retained and would require dedicated extraction and chromatographic methods. HRMS detection was achieved with a Q-Exactive hybrid quadrupole-Orbitrap (Thermo Fisher) equipped with an electrospray ionization source (ESI). HRMS1 spectra were acquired for masses ranging from m/z 100 to 1000, with a resolving power of 140,000 at m/z 200 and a mass error below 5 ppm. HRMS2 spectra were recorded in data-dependent mode with a resolving power of 17,500 (more details in SI section S-3.1). Separate analysis runs were conducted for positive and negative ionization modes with a spray voltage of +4 kV and −4 kV, respectively, and a capillary temperature of 300 °C. Confirmation of the prioritized structures was conducted at the University of Amsterdam. Reference standard materials and sample extracts were analyzed with an ultra-high-performance liquid chromatography (UHPLC) system (Nexera, Shimadzu, Den Bosch, The Netherlands) coupled to a maxis 4G q-ToF/HRMS equipped with an ESI source (Bruker Daltonics, Wormer, The Netherlands). Further details on this system are provided in SI section S-3.2).

Nontarget Screening Workflow. The NTS workflow consisted of three steps dealing with (i) HRMS1 data preprocessing, (ii) prioritization, and (iii) structure elucidation. Unless stated otherwise, the steps were automated and computed within R (v 3.3.2). For data preprocessing, the analyses raw files were converted to centroided mzXML format with ProteoWizard (v 3.0)44 and imported into enviMass (v 3.4). The enviMass settings used for this study are given in SI section S-4). Separate projects were created for positive and negative ESI data. Peak picking was performed to determine the nontarget features, i.e., unique m/z and retention time (t<sub>R</sub>) pairs. The 128 isotope-labeled ISs were used for mass recalibration, t<sub>R</sub> alignment, and intensity normalization; i.e., for each measurement, the intensities of the picked peaks were normalized using the median deviation of all ISs from their individual median profile intensity.40 The features were profiled; i.e., unique IDs were assigned to m/z and t<sub>R</sub> pairs detected across different samples. Features detected in analysis blanks, procedural blanks, and deionized water samples spiked with isotope-labeled ISs (termed blind samples) were profiled and subtracted. Profiled features that were not detected in all three replicates or whose t<sub>R</sub> was < 2 min and > 24 min were filtered out. The t<sub>R</sub> filter ensured that high polarity (t<sub>R</sub> < 2 min) and low polarity organics (t<sub>R</sub> > 24 min) were excluded from the data processing, as data of substances eluting either early or late are known to be of lesser quality with the existing chromatographic method, making assignment of (quasi-)isobaric substances challenging. For positive data, grouping of the most common single-charge ESI adducts, i.e., [M + H]<sup>+</sup>, [M + Na]<sup>+</sup>, [M + NH<sub>4</sub>]<sup>+</sup>, and [M + K]<sup>+</sup>, was additionally performed to define components.40 Other positive adducts, negative adducts, and isotopic peaks may increase the final number of profiled features, but were not included here.

The profiled features were prioritized by hierarchical cluster analysis (HCA) based on successful applications to lake sediments43 and to ozonation in a WWTP.42 Briefly, profile intensities were normalized to the maximum value detected in the whole data set, and dissimilarities expressed as Euclidean distance were calculated with the stats package (daisy function). The hierarchical classification of the profiled features was computed with the cluster package (hclust function). More details on the application of HCA to profile prioritization can be found elsewhere.33,34 The optimal number of clusters (k) was investigated by silhouette analysis computed with the cluster package (silhouette function). The average silhouette width, i.e., a dimensionless value indicating whether an object truly belongs to the cluster it was assigned to, was calculated at different k values.47 Once the optimal number of clusters was defined, the 15 most populated clusters were considered for further prioritization. In order to obtain good quality spectra for structure annotation, the top 50% most intense ions in each cluster were prioritized and their experimental HRMS1 and HRMS2 data extracted using the RMassBank package.48

Tentative structure elucidation was performed with MetFrag command line v 2.3 in batch mode.46 Neutral monoisotopic masses of the prioritized features were used to retrieve structures within 5 ppm mass accuracy from ChemSpider.31 The maximum number of candidates per feature was set to 5000. The maximum tree depth (M<sub>T</sub>) was set to 2. The candidate structures retrieved by MetFrag are fragmented in silico in a combinatorial manner and the fragments matched to the experimental HRMS2 spectra. Additionally, spectral similarities with records in the MoNA spectral library built into MetFrag were calculated using the MetFusion approach.49 All suspect lists available on the NORMAN Suspect List Exchange as of November 2016 were merged into one large list of 11,922 unique InChIKey codes. The candidate structures were scored based on seven terms: FragScore (in silico fragmentation score); MetFusionOffline (MoNA spectral similarities using the MetFusion approach); CSRefsScores (4 scoring terms: reference count on ChemSpider, reference count on PubMed, reference count on Royal Society of Chemistry and ChemSpider data sources); Suspects (hit in suspect list). The seven scoring terms were normalized by the highest value found among the proposed candidates, and an equal weighting of 1 was used to calculate the MetFrag Combined Score. An in-house script was used to check for agreement of formulas as calculated by MetFrag and the GenFormR package.50 Tentatively annotated spectra and extracted ion chromatograms (EICs) of the nontarget features were plotted with the packages ReSOLUTION51 and patRoon,52 respectively, for quality control. Finally, identification confidence levels were assigned to all the tentatively annotated features.29

### RESULTS AND DISCUSSION

#### Data Preprocessing, Prioritization, and Structure Annotation.

An overview on detection, mass deviation, and intensity ranges of the ISs is given in SI section S-5. Preprocessing of HRMS spectra with enviMass resulted in 10,850 positive and 7412 negative profiled features across the transect. The HCA results were visualized by plotting heat maps and dendrograms, shown in Figure 2.
The HCA classified intensity profiles based on similarities between the detected trends. The results revealed that the data in both polarities were characterized by a series of dynamic trends (clustered at the top of the heat maps) with a high-intensity region between wells LT-P09 and LT-P11, thus in water that originated from the river Lek throughout the 1990s. Additionally, each well seemed to display a set of unique features, visible on the heat maps as high-intensity spots. In the positive data, the spots displayed overall increasing dissimilarity with increasing RBF travel time (Figure 2a), whereas this behavior was not observed in the negative data (Figure 2b). While it could not be excluded that some compounds would occur only in one well, these profiles might result from detection above the peak-picking threshold value exclusively in one well or from lack of detection in all three replicates from adjacent wells. The optimal number of clusters ($k$) was determined by silhouette analysis. The average silhouette width values for $k = 80$ were 0.690 and 0.687 for positive and negative data, respectively, indicating reasonable separation of the data. The clusters were sorted by the number of profiled features and the 15 most populated clusters were inspected. Inspection of the prioritized clusters is discussed in the following section (see Interpretation of Clustered Time Trends).

The prioritized clusters were populated with 7525 positive and 5123 negative profiled features, of which 3764 and 3845 were prioritized based on their intensity. The MetFrag package was used to extract experimental HRMS1 and HRMS2 information, resulting in 1348 and 983 positive and negative profiled features, respectively, with associated fragmentation data. It should be noted that a higher percentage fraction of negative features was prioritized from the 15 most populated clusters on the basis of their intensities, as only about 25% of these ions triggered an HRMS2 acquisition. An overview of the population of each prioritized cluster, the number of features prioritized from these clusters and how many ions triggered...
HRMS2 acquisitions is shown in SI section S-6. Candidate structures were assigned to 884 and 550 positive and negative features, respectively, using MetFrag, whereas 369 positive and 345 negative features were excluded from further identification efforts as the maximum number of candidate structures was exceeded (set to 5000 to reduce runtime issues and eliminate cases with poor likelihood of success, based on experience). The MetFrag results were initially reduced by filtering out candidates with combined score lower than 4 (out of 7) and without any explained HRMS2 peaks. The minimum score filtering criterion proved too strict for negative data, which displayed lower scores overall, mostly due to lack of metadata, poor fit of the in silico fragments to the experimental data, or absence from the suspect list. Consequently, for negative data the minimum score was set to 3, including at least one explained HRMS2 peak. The MetFrag Scores plots were visualized for all tentatively annotated features to ease the assessment of the MetFrag results. An example of MetFrag Scores plot for the positive feature m/z 116.0165 detected at t_R 2.5 min is provided in Figure 3a.

The MetFrag Scores plot in Figure 3a shows individual and combined scores assigned to 59 candidates retrieved from ChemSpider and tentatively annotated to elucidate the non-target feature of interest, in this example, positive feature m/z 116.0165 (t_R = 2.5 min). The top 3 ranking structures were methylisothiazolinone, thiazole-5-methanol and 2-methoxy thiazole, which displayed a rounded score of 5.4, 4.4, and 3.8, respectively (CombScore, black line). The structure of methylisothiazolinone could explain 7/16 peaks, whereas the second and third candidates could explain 4/16 peaks (NoExplPeaks, red dashed line). The higher number of matching fragments resulted in a higher in silico fragmentation score for methylisothiazolinone (FragScore, red line). The top 3 candidates did not differ substantially in terms of spectral similarity scores (MetFusionOffline, blue dotted line), which was low overall and indicated a lack of similar or matching HRMS2 spectra in the library. The reference scores of the top 3 candidates were comparable, although higher for the first candidate (AvgCSRefs, green dot-dash line). The top 2 candidates also had a suspect hit (Suspects, purple dot).

Further processing involved generating a tentatively annotated HRMS2 spectrum from the output of MetFrag and GenFormR. MetFrag retrieved candidate structures from ChemSpider, generated fragments in silico, back-calculated their (de)protonated monoisotopic masses, and fitted them to the experimental HRMS2 data. GenFormR instead performed an algebraic calculation on spectral data to find the best formulas fitting the precursor and product ions. As these approaches are complementary, GenFormR was used to gain additional information to MetFrag to enhance the interpretation of the spectra. An example of tentatively annotated spectrum of the highest ranking candidate to elucidate the structure of positive feature m/z 116.0164 is shown in Figure 3b. The structure of methylisothiazolinone was eventually confirmed with a reference standard.

Applying the NTS workflow resulted in the tentative annotation of 72 nontarget features (all Level 3 at this point), 45 from positive clusters and 27 from negative clusters. The full list of tentatively annotated features, including final identification confidence levels, is given in SI section S-7. An overview of the data reduction at the different steps of the NTS workflow is shown in Figure 4.

Reference standards of 42 compounds were obtained based on availability and price. All chemicals were highest ranking candidates of their respective nontarget features, with the exception of atrazine-desethyl-2-hydroxy and metamitron-desamino, which were the second-highest ranking candidates, selected based on expert knowledge and following inspection of the MetFrag scores. Atrazine-desethyl-2-hydroxy was chosen for its higher in silico fragmentation score and spectral similarity score compared to the top candidate. The identification of metamitron-desamino represented an interesting case to demonstrate the importance of analyst judgment. This compound was the second-top ranking structure to explain positive ESI feature m/z 188.0817 at t_R 6.55 min with a MetFrag Combined Score of 3.25 and 11/17 explained HRMS2 peaks. The highest ranking structure for this feature was the phosphodiesterase 3 inhibitor amrinone, with a score of 5.74 and 10/17 annotated HRMS2 peaks. The most pronounced differences in the MetFrag scores of these two structures were found in the number of references. Since one of the scoring criteria was the number of PubMed references, it was not surprising that the pharmaceutical amrinone had a higher score than a pesticide TP. However, since nearly all tentatively annotated chemicals originated from either industrial or agricultural activities, metamitron-desamino was thought to be the more likely structure. Reference standards were obtained for both compounds, leading to confirmation of the pesticide TP. It is noteworthy that metamitron-desamino would have been missed without amrinone being the first candidate, as its MetFrag score was below the cutoff value used for prioritization of the MetFrag results of the positive ESI data.
Methylisothiazolinone C₆H₇NOS 115.009 [M + H]+ 2.5 2.5 36393
1,3-Benzothiazole C₆H₇NS 135.014 [M + H]+ 7.5 7.5 6952
Tetramethylexylsulfamide C₆H₁₀N₂O₅S 152.062 [M + H]+ 4.8 4.8 121689
Atrazine-desethyl-2-hydroxy C₆H₁₀N₂O 169.096 [M + H]+ 3.1 3.0 96906
4-Toluenesulfonamide C₆H₇NO₃S 187.035 [M + H]+ 5.6 5.6 6033
Simazine-2-hydroxy C₆H₇NO₃S 183.112 [M + H]+ 4.6 4.5 16505
Metamitron-desamino C₈H₈N₄O₃S 187.075 [M + H]+ 6.0 6.0 157884
Benzoguanamine C₆H₇N₅ 187.086 [M + H]+ 5.0 4.9 6797
2,6-dichlorobenzamide C₇H₈ClN₂O 188.975 [M + H]+ 5.0 5.0 15359
Carbendazim C₇H₈N₂O₂S 191.069 [M + H]+ 6.0 6.0 157884
Chlorotoluron C₇H₇ClN₅O 212.072 [M + H]+ 9.1 9.1 25472
Dichlorodimethyl C₈H₇Cl₂N₂O 218.001 [M + H]+ 9.8 9.8 18040
Diphenylphosphinic acid C₉H₇ClNO₃S 223.030 [M + H]+ 6.6 6.6 8172
5-Amino-2-chlorotoluene-4-sulfonic acid C₇H₇ClNO₃S 220.991 [M + H]+ 5.4 5.4 6670
Chloridazon C₇H₇ClNO₃S 221.036 [M + H]+ 6.0 6.0 14790
Naphthonic acid C₇H₇ClNO₃S 223.030 [M + H]+ 5.4 5.4 6670
Lamotrigine C₉H₇ClN₅ 255.008 [M + H]+ 6.7 6.6 3741
Trityl phosphate C₉H₇ClN₃O₃P 266.165 [M + H]+ 14.9 15.0 29090
Acesulfame C₉H₉ClNO₄S 269.094 [M – H]− 2.5 2.5 33607
O₂O-diethyl thiophosphate C₉H₈N₂0₆S 170.017 [M – H]− 3.6 3.6 635
p-Toluidine-n-sulfonic acid C₇H₇N₂O₃S 187.030 [M – H]− 3.3 3.3 60405
Dibutyl phosphate C₁₀H₂₃O₄P 210.102 [M – H]− 8.8 8.8 7593
Camphorsulfonic acid C₁₀H₁₂O₄P 232.077 [M – H]− 5.0 5.0 17438
4-Amino-2,5-dichlorobenzenesulfonic acid C₉H₈ClNO₃S 240.937 [M – H]− 2.6 2.6 59986
4-Dodecylbenzenesulfonic acid C₁₂H₂₃O₄S 326.192 [M – H]− 17.3 17.3 8172

Table 1. List of Nontarget Contaminants Confirmed with Reference Standards, Their Formula, Neutral Monoisotopic Mass, Detected Adduct, Retention Time (tᵣ) in Samples and Standards, and ChemSpider ID

The candidate structures of 25 out of 42 nontarget features were confirmed, resulting in a success rate of 60% against the confirmation subset. An extensive evaluation of the performance of MetFrag can be found in the literature.56,37 Details of the successfully confirmed compounds are shown in Table 1. Of the 17 compounds that were not confirmed, 4 could not be detected by the UHPLC-ESI-q-TOF/HRMS system and 13 eluted at a different tᵣ.

Interpretation of Clustered Time Trends. Several prioritized clusters displayed trends with intensity maxima in the 1990s (wells LT-P09 and LT-P11); there were seven in the positive ESI data and six in the ESI negative data. The presence of these clusters could be rationalized by the history of anthropogenic emissions and restoration measures in the Rhine basin from the 1950s onward. River protection programs coordinated by the International Commission for the Protection of the Rhine (ICPR) resulted in domestic and industrial WWTPs being built between 1970 and 1990 in the Rhine riparian countries.53 A turning point in international river basin management was the 1986 Sandoz accident when 15,000 m³ of water mixed with organophosphorus and organochlorine compounds including pesticides, dyes, solvents, and intermediates accidentally entered the upper Rhine in Switzerland, resulting in widespread mortality of fish, macroinvertebrate, and plankton communities in the riparian countries downstream.53–56 Following the Sandoz accident, the ICPR established the Rhine Action Programme (RAP) to coordinate and implement measures to lower the discharge of hazardous substances. Their goals were achieved in the mid-1990s when substantial reduction of organic emissions, improved oxygen content, and increased biodiversity were reported.53

Given the high number of (tentatively) identified substances, the discussion of the results is limited to the most significant findings from the confirmed identities, whose profiles are highlighted in Figure 5. The cluster shown in Figure 5a indicated a net increase of emissions from the 1950s to the 1990s, followed by a moderate but constant decrease in the 2000s. In this cluster, a negligible number of profiles displayed intensity maxima in the early 2000s; however, these are not discussed further, as no candidate structures could be associated with these trends. Based on the known sources and environmental fate of 2,6-dichlorobenzamide56 (Figure 5a—green line), nonpoint source emissions from agricultural applications of biocides or sewage farming could explain some of these profiles.

Confirming the environmental occurrence of tetramethylsulfamide (TMS) was a novel discovery of the present study (Figure 5a—blue line), highlighting the environmental significance of NTS. No information about production or use of TMS was found on the European Chemical Agency (ECHA) or the US EPA Web sites. TMS was not present in the NORMAN Suspect Lists Exchange or in accurate mass spectral libraries. A search on PubChem returned 50 patents detailing formulations for the synthesis of dyes, flame retardants, and pesticides. TMS can be used for the synthesis of sulfur trioxide-dimethylamine complex, a sulfating reagent for dyestuffs.56 TMS can also be a byproduct of the synthesis of dimethylsulfamyl chloride, a chemical used in a variety of industrial applications,58 or a byproduct of the synthesis of sulfur-containing aziridines chemosterilants.59 The exact
The identity of 4-toluenesulfonamide (4-TSA) in a cluster with maxima in the 1990s was confirmed with a reference standard (Figure S5a—red line). 4-TSA is a plasticizer, an intermediate for the synthesis of pesticides and disinfection byproduct of the antimicrobial agent N-sodium-N-chloro-p-toluenesulfonamide. 4-TSA was detected in the Berlin area in surface water, groundwater, and bank filtrate at concentrations up to 0.9 μg/L, 14.1 μg/L, and 0.24 ng/L, respectively. The authors of that study pointed to untreated WWTP effluents and former sewage farms as sources of sulfonamides and concluded that 4-TSA can help identify bank filtrate originating from polluted surface water, supporting the results of the present work.

A more dynamic intensity trend showing increasing emission from the 1950s to the 1990s followed by a substantial decrease in the 2000s is shown in Figure 5b. In this cluster, camphorsulfonic acid (CSA) was tentatively identified and later confirmed, along with the metabolite O,O-diethyl thiophosphate (DETP), the industrial chemicals 5-amino-2-chlorotoluene-4-sulfonic acid (ADCBSA). To the best of our knowledge, this is the first time CSA has been confirmed in bank filtrate and its environmental persistence demonstrated (Figure 5b—blue line), highlighting the environmental significance of NTS. CSA was not included in the suspect list, and only one reference was found reporting its occurrence in a WWTP effluent from a rubber manufacturing site in Spain. CSA is used as a dopant in the synthesis of polyaniline, a conductive polymer. Camphor derivatives, such as CSA, are used as UV filters in cosmetic products and eventually reach surface waters via insufficiently treated domestic WWTP effluents. The UV filter terephthalidene dicamphorsulfonic acid (TPDCSA) is unstable under photolysis in aqueous media and uncharacterized degradation product(s) with UV absorbance <290 nm were reported. CSA has a UV/vis absorbance of 285 nm, suggesting that it may originate from TPDCSA in the environment, rather than exclusively from industrial sources.

ACTSA (Figure 5b—gray line) is an important building block for the synthesis of dyes reported to persist in chemical and biological wastewater treatment. Its 2-amino isomer can be obtained from the cleavage of the azo dye Pigment Red 53 and was included in a priority list of 23 unregulated aromatic amines of toxicological concern. Limited literature references were found for this compound, and to the best of our knowledge, this is the first time it has been identified in a riverbank filtrate. The REACH registration dossier of ACTSA was submitted by a dye manufacturer with a production site at the river Main, the longest tributary of the Rhine, approximately 500 km upstream of the RBF site. The identity...
of another dyestuff synthesis intermediate, ADCBSA, was confirmed in this cluster (Figure 5b—green line). This chemical occurs in liquid waste from manufacturing processes; however, liquid waste containing ADCBSA is usually treated separately before being sent to WWTPs, because this compound is toxic to microorganisms and inhibits biological treatment. Even if nonpoint sources cannot be excluded, the stability trend from the mid-1990s onward. The lower introduced in the European Union in 1975 and has displayed a sales data of its parent compound metamitron, which was from rivers impacted by urban and agricultural activities. Metamitron-desamino up to 680 ng/L have been reported potential to leach into groundwater. Concentrations of displayed stable intensities along the transect were found is the main biodegradation product of the herbicide 5c, the metabolites metamitron-desamino, atrazine-desethyl-2-hydroxy and atrazine-desethyl-2-hydroxy showed persistence across the transect (Figure 5d—blue line). Previous studies found that simazine occurred in surface water and groundwater at concentrations up to 10 times lower than atrazine. It is unclear whether simazine-2-hydroxy was removed during RBF, transformed further or diluted to undetectable concentrations in the older bank filtrate. Both triazine herbicides that lead to formation of these TPs were banned in the European Union in 2004. The detection of atrazine-desethyl-2-hydroxy at constant levels might be attributed to the release of atrazine (or TPs) from contaminated river sediments prior to infiltration in the RBF system. Other known metabolites of atrazine and simazine were screened for in HRMS1 data, but were not detected. These TPs might have either been absent from the RBF transect or occurred at undetectable concentrations. It cannot be excluded that atrazine-desethyl-2-hydroxy prevailed over other more commonly detected dealkylated metabolites, as the aquifer screened in this study differs from others for being both confined and anaerobic. For example, in a recent screening of three aerobic bank filtration sites, atrazine-2-hydroxy and atrazine-desethyl were both found in low ng/L concentration, whereas atrazine-desethyl-2-hydroxy and atrazine-desisopropyl were not detected.

### Environmental Implications

This study contributes to the mounting evidence of environmental persistence of hydrophilic organic compounds and shows that polar substances can be highly mobile in RBF systems with long travel time at stable biogeochemical conditions. More research should be done on this or comparable RBF transects to investigate the fate of the most polar MPs, e.g., with even lower logD values than those
identified in this study, which might have been lost during enrichment or insufficiently retained by reversed-phase chromatography. We showed that state-of-the-art NTS relying on open computational tools and performed in a semi-automated manner can be an extremely powerful method to explore water contaminants by HRMS. Spectra of the substances identified in this study will be uploaded to openly accessible accurate mass spectral libraries to contribute to future screenings. The list of confirmed contaminants has been shared with local drinking water utilities to assess their removal in drinking water treatment. The trend analysis presented here can be useful to manage bank filtration systems with long travel times in catchment areas impacted by anthropogenic activities. In these cases, to avoid contamination with many legacy pollutants from the 1990s, which overall displayed higher normalized intensities in the older water and thus likely occurred at higher concentrations, the use of young groundwater to produce potable water is recommended, whereas advanced treatment should be applied to the old groundwater. In the case of drinking water treatment at the Tiendweg well field, reverse osmosis is applied to maximize micropollutant removal.

■ ASSOCIATED CONTENT

2 Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.est.9b01750.

Standards and chemicals and their sources; tables of labeled and unlabeled standards with molecular formulas and CAS numbers. Additional details on the study area; figure showing the well fields with groundwater flow lines and modeled travel times. Additional details on the HRMS systems. Settings for enviMass. Overview on the detection of the isotope-labeled standards; figures showing mass deviation and intensity distribution of the labeled-standards; tables listing detected adducts, retention time and sample matches of the isotope-labeled standards. Tables listing feature population of the prioritized clusters, number of prioritized features and features with associated HRMS2 data. List of (tentatively) identified substances. (PDF)

■ REFERENCES


Filtration Sites to Assess Natural Attenuation and Potential pollutant Screening Using LC-HRMS/MS at Three Riverbank...

Huggenberger, P.; Borer, P.; Singer, H. Comprehensive Micro-


Hollender, J; Rothardt, J; Radny, D.; Loos, M.; Epting, J; Huggenberger, P.; Borer, P; Singer, H. Comprehensive Micro-


Hollender, J; Rothardt, J; Radny, D.; Loos, M.; Epting, J; Huggenberger, P.; Borer, P; Singer, H. Comprehensive Micro-

Hollender, J; Rothardt, J; Radny, D.; Loos, M.; Epting, J; Huggenberger, P.; Borer, P; Singer, H. Comprehensive Micro-

Hollender, J; Rothardt, J; Radny, D.; Loos, M.; Epting, J; Huggenberger, P.; Borer, P; Singer, H. Comprehensive Micro-

Hollender, J; Rothardt, J; Radny, D.; Loos, M.; Epting, J; Huggenberger, P.; Borer, P; Singer, H. Comprehensive Micro-


